Catch me if you can: a spatial model for a brake-driven gene drive reversal

Vincent Calvez, Florence Débarre, Léo Girardin

▶ **To cite this version:**

Vincent Calvez, Florence Débarre, Léo Girardin. Catch me if you can: a spatial model for a brake-driven gene drive reversal. 2018. hal-01954463v1

HAL Id: hal-01954463

https://hal.science/hal-01954463v1

Preprint submitted on 14 Dec 2018 (v1), last revised 11 Jul 2019 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CATCH ME IF YOU CAN: A SPATIAL MODEL FOR A BRAKE-DRIVEN GENE DRIVE REVERSAL

VINCENT CALVEZ, FLORENCE DÉBARRE, AND LÉO GIRARDIN

Abstract. We consider a reaction–diffusion system modeling the co-invasion of a gene drive (an allele biasing inheritance, increasing its own transmission to offspring) and a brake (an "antidote" to a gene drive) in a population carrying a wild-type allele. We successfully prove that, whenever the drive fitness is at most 50% of the wild-type one while the brake fitness is close to the wild-type one, co-extinction of the brake and the drive occurs in the long run. On the contrary, if the drive fitness is at least 50% of the wild-type one, then co-extinction is impossible. Based on numerical experiments, we argue in favor of a global co-extinction conjecture provided the drive fitness is at most 50% of the wild-type one. The proof relies upon the study of a related predator–prey system with strong Allee effect on the prey.

1. Introduction

With the development of CRISPR-Cas9 genetic engineering, population management using gene drives has become a realistic possibility. The technique consists in artificially biasing the inheritance of a trait of interest in a target population [5]. Such biased inheritance is due to the presence of an artificial self-replicating element expressing a DNA-cutting enzyme (the Cas9 endonuclease), such that initially heterozygous individuals (*i.e.* carrying the drive construct on one chromosome and a wild-type sequence on the homologous chromosome) produce almost exclusively drive-carrying gametes instead of 50% of drive-carrying gametes, as expected under Mendelian segregation (see Figure 1.1). The proportion of drive-carrying gametes depends on the conversion efficiency, perfect conversion meaning that an initially heterozygous individual produces 100% of drive-carrying gametes. Thanks to its supra-Mendelian transmission to offspring, a drive can spread in a population even if it confers a significant fitness cost [11]. Potential applications for human health and agriculture include the modification of mosquito populations to make them resistant to malaria or the eradication of agricultural pest species [8].

(V. C.) Institut Camille Jordan, UMR 5208 CNRS & Université Claude Bernard Lyon 1, France

(F. D.) CNRS, Sorbonne Université, Université Paris Est Créteil, Université Paris Diderot, INRA, IRD, Institute of Ecology and Environmental sciences - Paris, IEES-Paris, 75005 Paris, France

(L. G.) Laboratoire de Mathématiques d'Orsay, Université Paris Sud, CNRS, Université Paris-Saclay, 91405 Orsay Cedex, France

E-mail addresses: vincent.calvez@math.cnrs.fr, florence.debarre@normalesup.org, leo.girardin@math.u-psud.fr.

2000 *Mathematics Subject Classification.* 35K57, 37N25, 92D10, 92D25.

Key words and phrases. long-time behavior, gene drive, brake, predator-prey, strong Allee effect.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 639638). This work was supported by a public grant as part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH, and ANR-14-ACHN-0003-01.

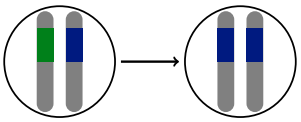


FIGURE 1.1. Perfect gene conversion: a heterozygous individual only produces gametes carrying the gene drive allele.

However, while very promising, the technique is not risk-free. A drive could have off-target effects, or spread in a non-target population. For instance, a drive can be introduced on an island to eradicate a local rat population, but the dispersal of drive-carrying individuals to the mainland or to another island would also threaten those populations [4]. The effects of population modification using a drive may also have unexpected consequences on other species, *e.g.*, predators or competitors. More generally, it is important to be able to control the spread of a drive, and to stop it if necessary. To this end, a "brake" construct was proposed, that does not contain the *cas9* gene (and is hence unable to convert a wild-type allele), but that is able to target the very *cas9* sequence contained in a drive construct, and therefore to convert a drive allele into a brake allele in a (drive/brake) heterozygote [16]. The construct has been shown to not only stop a drive, but also in some cases lead to the recovery of the original wild-type population [13].

Most models of gene drive consider well-mixed populations (except meta-population and partial differential equations models, [2, 9, 11]). Here we consider the influence of space and limited dispersal, and ask whether a brake construct is able to stop the spatial spread of a drive. More precisely, we introduce a minimalist PDE model of two interacting sub-populations combining spatial diffusion and Mendelian population genetics. The two sub-populations refer to the frequency of individuals carrying respectively the gene drive allele and the brake allele embedded in a wild-type population. The case of a single sub-population of gene drive was addressed in [11]. It was shown that the spatial invasion of the gene drive allele was successful up to a level of approximately 70% of fitness cost, see below for details. In the present work, we ask whether the brake, even if lately introduced (with a spatial delay with respect to the gene drive invasion) can catch up the invasion and hinder the gene drive propagation. Using phase plane analysis and techniques from reaction–diffusion PDE, we give a positive answer to this question, under some conditions on the model parameters. We even prove that the gene drive frequency may be eventually reduced to zero everywhere, provided that the fitness cost of the drive is above 50%, and that the fitness cost of the brake is small enough.

1.1. **The model.** We adopt the following set of notations:

- $N \in \mathbb{N}$ is the spatial dimension (typically $N \in \{1, 2, 3\}$);
- $n \in [0, +\infty)$ is the total population density;
- $u \in [0, 1]$ and $v \in [0, 1]$ $(0 \leq u + v \leq 1)$ are the respective frequencies in the population n of the gene drive allele D and the brake allele B (so that the frequency of the wild-type allele O is exactly $1 - u - v$);
- $a \in (0, 1)$ and $b \in (0, 1)$ are the respective selective disadvantage (*i.e.*, decreased survival) of the homozygous individuals DD and BB compared to the wild-type, with the assumption $a \geq b$, which is biologically relevant[1];

[1]Depending on the construct, the brake could just convert a drive without affecting its effect on fitness (b close to a, $b \leq a$), or at the other extreme the brake could carry a cargo gene restoring wild-type fitness (b close to 0).

- $h \in [0,1]$ is the dominance of the brake allele B on the wild-type allele O (in particular, B is dominant if $h = 1$, recessive if $h = 0$ and additive if $h = 1/2$).

Following the biological literature [11, 12], we assume that:

(1) the conversion efficiency of the drive (conversion $OD \to DD$) and of the brake (conversion $DB \to BB$) are perfect[2];
(2) gene conversion takes place early in development (*e.g.*, an individual born as OD becomes DD and has the fitness of a DD individual)[3];
(3) gametes mate randomly (*i.e.* according to a uniform law).

Then the spatially homogeneous, next-generation discrete system reads:

$$\begin{cases} 1 - u_{g+1} - v_{g+1} = \dfrac{(1-u_g-v_g)\left(1-u_g-v_g+(1-hb)\,v_g\right)}{1-\left(au_g^2+bv_g^2+2bu_gv_g\right)-2\,(1-u_g-v_g)\,(au_g+hbv_g)}, \\ u_{g+1} = \dfrac{u_g\left((1-a)\,u_g+2\,(1-a)\,(1-u_g-v_g)\right)}{1-\left(au_g^2+bv_g^2+2bu_gv_g\right)-2\,(1-u_g-v_g)\,(au_g+hbv_g)}, \\ v_{g+1} = \dfrac{v_g\left((1-b)\,v_g+2\,(1-b)\,u_g+(1-hb)\,(1-u_g-v_g)\right)}{1-\left(au_g^2+bv_g^2+2bu_gv_g\right)-2\,(1-u_g-v_g)\,(au_g+hbv_g)}. \end{cases}$$

The first line corresponds to the frequency dynamics of the wild-type allele, the second line to the drive, and the last line to the brake. For each equation, the numerator on the right-hand side corresponds to the amount of corresponding alleles produced, while the denominator corresponds to the total amount of alleles (*i.e.*, the "mean fitness" in the population). It can be verified that this denominator is exactly such that the first equation is true (in other words, such that the sum of the three frequencies remains identically equal to 1 as time goes on). Since the first equation is redundant, we get rid of it hereafter.

These equations can be understood as follows:

- Wild-type alleles (first equation) are carried *(i)* by all of the gametes produced by OO homozygotes (initially in frequency $(1-u_g-v_g)^2$, and with fitness 1), and *(ii)* by half of the gametes of OB heterozygotes (in frequency $2v_g(1-u_g-v_g)$, and with fitness $(1-hb)$). Note that since gene conversion is assumed to be perfect, no wild-type alleles are produced by initially OD individuals.
- Drive alleles (second equation) are carried *(i)* by all of the gametes produced by DD homozygotes (initially in frequency u_g^2 and who have a fitness $(1-a)$), but also *(ii)* by all of the gametes produced by initially OD heterozygotes (initially in frequency $2u_g(1-u_g-v_g)$), who were immediately converted into DD homozygotes, and hence have fitness $(1-a)$. Since gene conversion is assumed to be perfect, no drive alleles are produced by initially DB individuals.
- Brake alleles (third line) are carried *(i)* by all of the BB homozygotes (initially in frequency v_g^2, and who have a fitness $(1-b)$), *(ii)* by all of the gametes produced by initially DB heterozygotes (in frequency $2u_gv_g$), who were immediately converted into BB homozygotes, and hence have fitness $(1-b)$, and finally *(iii)* by half of the gametes produced by OB heterozygotes (in frequency $2v_g(1-u_g-v_g)$, and with fitness $(1-hb)$).

Next, we take the spatial diffusion of the individuals of the population n into account and assume:

[2]This is merely for algebraic convenience and the general case will be discussed below in Subsection 4.2.

[3]With late gene conversion (typically in the gonads), an OD-born individual would have the fitness of an OD. In both cases though, only D gametes are produced by this individual.

(4) the time scale of the diffusion mechanism and the maturation time between two generations are of the same order;
(5) each sub-population diffuses at the same rate.

We are now in position to perform a classical first-order approximation and obtain the following reaction–diffusion PDE system:

$$\begin{cases} \dfrac{\partial u}{\partial t} - \Delta u - 2\nabla(\log n)\cdot\nabla u = u\left(\dfrac{(1-a)\,u + 2\,(1-a)\,(1-u-v)}{1-(au^2+bv^2+2buv)-2\,(1-u-v)\,(au+hbv)} - 1\right), \\ \dfrac{\partial v}{\partial t} - \Delta v - 2\nabla(\log n)\cdot\nabla v = v\left(\dfrac{(1-b)\,v + 2\,(1-b)\,u + (1-hb)\,(1-u-v)}{1-(au^2+bv^2+2buv)-2\,(1-u-v)\,(au+hbv)} - 1\right). \end{cases}$$

The transport term $2\nabla(\log n)$ is the signature of gene flow in populations which are not homogeneous in size. This is an easy consequence of the reformulation of the problem by means of frequencies rather than population densities. To make this connection clear, we point out that the diffusion operator $\frac{\partial}{\partial t} - \Delta$ for the population density un is related to the diffusion–transport operator $\frac{\partial}{\partial t} - \Delta - 2\nabla(\log n)$ for the frequency u:

$$\begin{aligned} \frac{\partial\,(un)}{\partial t} - \Delta\,(un) &= \frac{\partial u}{\partial t}n + \frac{\partial n}{\partial t}u - (\Delta u)\,n - (\Delta n)\,u - 2\nabla u\cdot\nabla n \\ &= n\left[\frac{\partial u}{\partial t} - \Delta u - 2\nabla(\log n)\cdot\nabla u\right] - u\left[\frac{\partial n}{\partial t} - \Delta n\right]. \end{aligned}$$

We shall address two types of questions in our discussion:

(1) what is the transient behaviour of each sub-population?
(2) what is the final outcome of the model?

In the latter, we can handle the *a priori* unknown transport term $2\nabla(\log n)$, under some mild assumption. However, we were not able to handle this transport term to discuss the former question, simply because this transport term may have an impact on the propagation speeds, although it should not modify the relative speed. This is the reason why we may sometimes restrict to homogeneous populations in some parts of the discussion, although we are aware that it is a serious restriction.

For well-posedness purposes, we assume that n is positive everywhere in time and space. It might asymptotically vanish but we further require some minimal regularity: we assume that $\nabla(\log n)$ is uniformly Hölder-continuous with an exponent larger than $\frac{1}{2}$ (for instance, uniformly Lipschitz-continuous). These mathematical assumptions should not be understood as being too restrictive: they are satisfied, for instance, in a large class of reaction-diffusion problems with initial conditions bounded above and below by positive constants.

For ease of reading, we define the scalar parabolic operator

$$\mathscr{P} = \frac{\partial}{\partial t} - \Delta - 2\nabla(\log n)\cdot\nabla$$

as well as the reaction functions

$$\begin{aligned} w\,(u,v) &= 1 - \left(au^2+bv^2+2buv\right) - 2\,(1-u-v)\,(au+hbv)\,, \\ g\,(u,v) &= \begin{pmatrix} (1-a)\,u + 2\,(1-a)\,(1-u-v) \\ (1-b)\,v + 2\,(1-b)\,u + (1-hb)\,(1-u-v) \end{pmatrix}, \\ f\,(u,v) &= \frac{1}{w\,(u,v)}\,g\,(u,v) - \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \end{aligned}$$

so that the system finally reads:

$$\mathscr{P}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} \circ f\,(u,v)\,. \tag{1.1}$$

Here, the fact that the 2×2 linear parabolic operator on the left-hand side is the same on both lines is of the utmost importance. Mathematically, it is a necessary and sufficient condition to apply several theorems of the standard parabolic theory, in particular a generalized maximum principle due to Weinberger [15] that we will indeed use extensively. Biologically, it means that the individuals carrying the allele D and those carrying the allele B move in space similarly: the gene under consideration does not affect the motility of the individuals carrying it.

Since u and v are frequencies, they satisfy $u+v\leq 1$ (with $1-(u+v)$ the frequency of the wild-type allele O). Therefore it is natural to define the triangle

$$\mathsf{T}=\left\{(u,v)\in[0,1]^2 \;\middle|\; u+v\leq 1\right\}.$$

It is of course the maximal closed invariant subset of the phase plane we are interested in (the invariance will be verified rigorously later on).

1.2. **Results.**

1.2.1. *Coexistence.* Our first result deals with the case of coexistence. For technical reasons that will be discussed below, the statement is restricted to spatially homogeneous solutions, *i.e.* we consider the solutions of the system of ODE with the righ-hand-side of (1.1) only.

In the case where the selective disadvantage of the drive compared to the wild-type is less than a half ($a<1/2$), we establish that both sub-populations of drive and brake persist in the long term, as measured by their frequencies which are positive at arbitrary large times. The mathematical statement is as follows.

Theorem 1.1. *Let (u,v) be a spatially homogeneous solution of (1.1) with initial condition in the interior of T.*

If $a<\frac{1}{2}$, then co-extinction of (u,v) cannot occur. More precisely,

$$\limsup_{t\to+\infty}\left(u\left(t\right)+v\left(t\right)\right)>0.$$

In view of numerical experiments, we expect damped oscillations converging to a coexistence state or sustained oscillations, periodic or approaching the concatenation of three heteroclinic orbits (see Subsection 2.3 below).

Concerning the spatially heterogeneous problem, describing the invasion of the drive and the brake in a territory occupied by the wild-type allele, we explain in Subsection 4.1 why we expect that the brake always catches up with the drive and that, afterwards, both populations persist in the wake of the joint invasion front. We give evidence that this claim is true if $a\leq 1/4$ and the total population n is spatially homogeneous, based on Ducrot–Giletti–Matano [3]. We believe it remains true even if $1/4<a<1/2$. Figure 1.2 is a numerical illustration of this claim. Indeed, we observe that the brake catches up with the drive, even if it starts with a space and time delay. However, the drive is "strong" enough to persist, resulting in a joint invasion front followed by a complicated spatio-temporal pattern.

1.2.2. *Coextinction.* Our second result deals with the case of coextinction. Here we are able to handle spatially heterogeneous solutions.

We assume that the initial frequencies (u_0,v_0) are distributed such that some individuals carrying the brake allele have been released somewhere, whereas the gene drive allele has not completed its invasion in the whole space yet (only individuals carrying the wild-type allele are present far away in space). If the drive pays a sufficient fitness cost ($a>1/2$), and if the selective disadvantage of the brake is not too large ($b<b^*(a)$), then the drive goes extinct everywhere, followed by the complete extinction of the brake as well. The threshold for the brake $b^*(a)$ appears for technical reasons at several steps of our argument. Consequently,

an explicit value is not straightforward to obtain, and might not be informative. We believe that this is only a technical restriction, and that the result should remain true for relatively large values of b. The mathematical statement is as follows.

Theorem 1.2. *Let (u,v) be the solution of (1.1) with initial condition $(u_0,v_0)\in\mathscr{C}\left(\mathbb{R}^N,\mathsf{T}\right)$ satisfying*

$$v_0\neq 0 \quad \textit{and} \quad \lim_{\|x\|\to+\infty}(u_0,v_0)(x)=(0,0).$$

There exists $b^\star(a)\in(0,a)$ such that, if $a>\frac{1}{2}$ and $b<b^\star(a)$, then co-extinction occurs:

$$\lim_{t\to+\infty}\left(\sup_{x\in\mathbb{R}^N}u(t,x)+\sup_{x\in\mathbb{R}^N}v(t,x)\right)=0.$$

It is important to notice that the convergence to zero of both u and v is uniform in space, that is, we rule out the possibility of a persistent wave of u followed or replaced by a wave in v.

We performed numerical experiments to explore the possible behaviours for relatively large values of b. We observed complete extinction for values of b up to a (Figure 1.3), and even beyond (Figure 1.4(A)). However, if the selective disadvantage of the brake is too large, then it does not succeed in blocking the invasion of the drive. Nevertheless, the case $b>a$ is beyond the scope of our assumptions and does not correspond to an interesting case in the context of the biological problem.

To complete the numerical investigation, we give some evidence in Subsection 4.1 supporting the conjecture that the co-extinction of u and v occurs for all $b\leq a$.

1.3. **Structure of the paper.** Section 2 is devoted to some technical preliminaries and contains in particular an elementary proof of Theorem 1.1. Section 3 contains the proof of Theorem 1.2. Section 4 is where possible extensions are discussed and where the aforementioned conjecture is explained.

2. Preliminaries

2.1. Well-posedness.

Proposition 2.1. *The function w satisfies*

$$\max_{\mathsf{T}} w=1 \ \textit{and}\ \min_{\mathsf{T}} w=1-a>0.$$

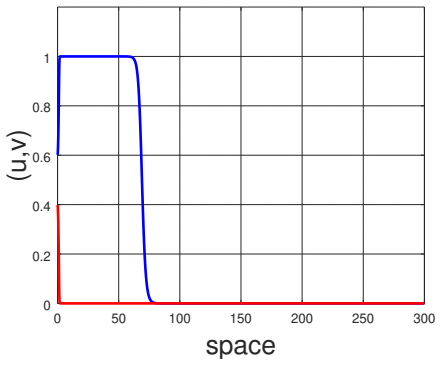
Therefore the system (1.1) is well-posed in T.

Proof. The function $u\mapsto w(u,0)$ coincides with $u\mapsto 1-au^2-2au(1-u)$, that is $u\mapsto 1-a+a(1-u)^2$, whose minimum and maximum in $[0,1]$ are respectively $1-a>0$ and 1. Since the bound $w\leq 1$ in T is obvious, this directly shows that the maximum of w in T is indeed 1. It only remains to confirm that the minimum is indeed $1-a$. We prove this claim in two steps: first, we prove that the minimum of w on $\partial\mathsf{T}$ is $1-a$; second, we prove that there is no critical point of w in the interior of T.

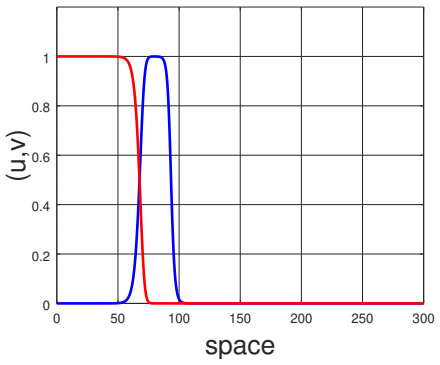
The function $v\mapsto w(0,v)$ coincides with $v\mapsto 1-bv^2-2hbv(1-v)$, that is $v\mapsto 1+(2h-1)bv^2-2hbv$. Its derivative with respect to b is $2b\{v(h-1)+h(v-1)\}\leq 0$. Therefore, its minimum in $[0,1]$ is attained at $v=1$ and is $1-b>0$.

The function $u\mapsto w(u,1-u)$ coincides with $u\mapsto 1-au^2-b(1-u)^2-2bu(1-u)$, that is $u\mapsto 1-b-(a-b)u^2$, whose minimum in $[0,1]$ is $1-a>0$.

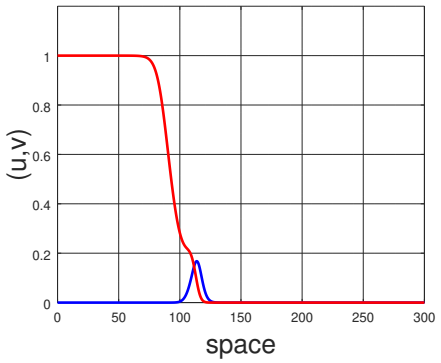
Therefore the minimum of w on $\partial\mathsf{T}$ is indeed $1-a>0$.



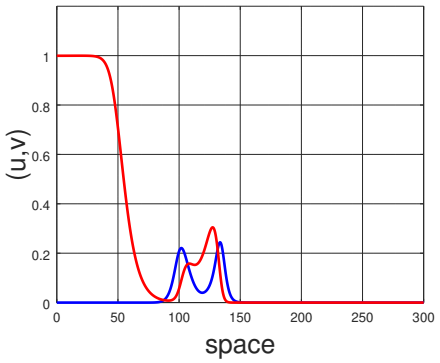
(A) $t = 80$



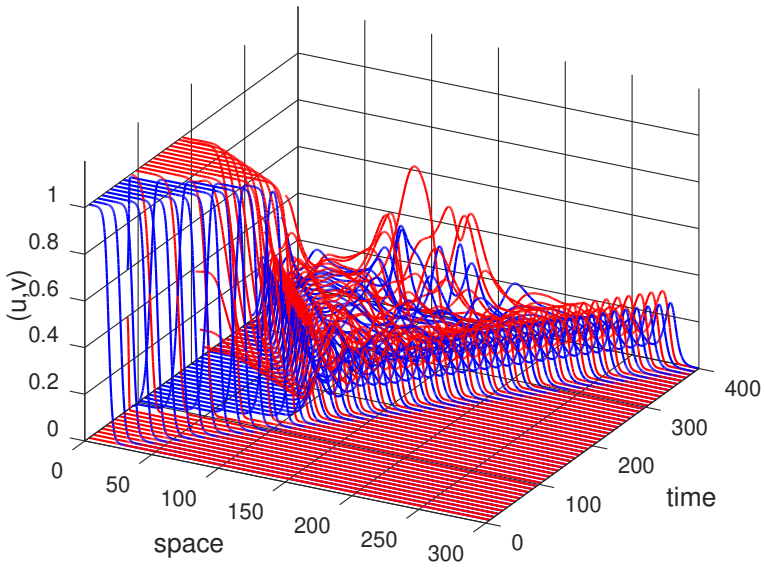
(B) $t = 112$



(C) $t = 144$

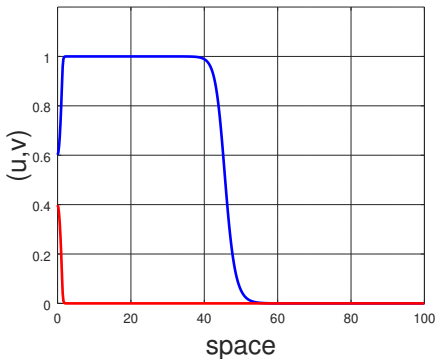


(D) $t = 176$

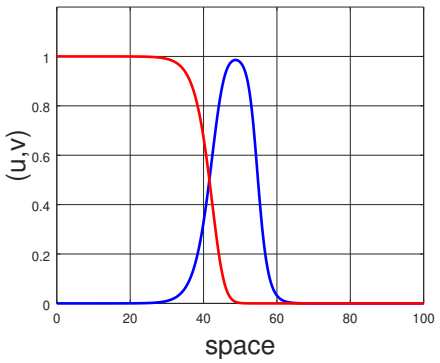


(E) Space-time representation of the solution (u, v)

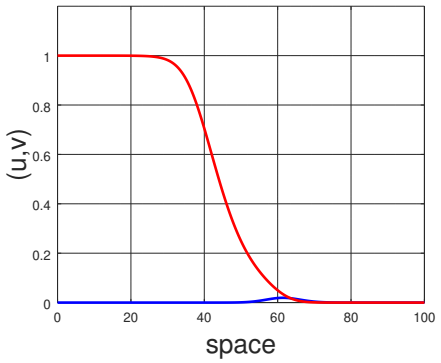
FIGURE 1.2. Numerical solution of (1.1) in the case $a = 0.45 < 1/2$, and $b = 0.35 < a$, $h = 0.5$, and n is spatially homogeneous (*i.e.* the drift term $\nabla \log n$ vanishes). (A-B-C-D) are successive snapshots of the drive u (blue) and the brake v (red) from the time $t = 80$ at which the brake v is released on the left-hand-side of the domain. (E) is the superposition of many snapshots, in order to visualize the spatio-temporal dynamics.



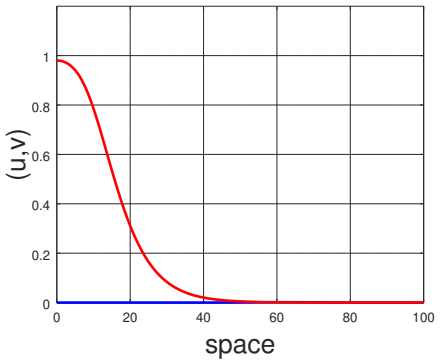
(A) $t = 80$



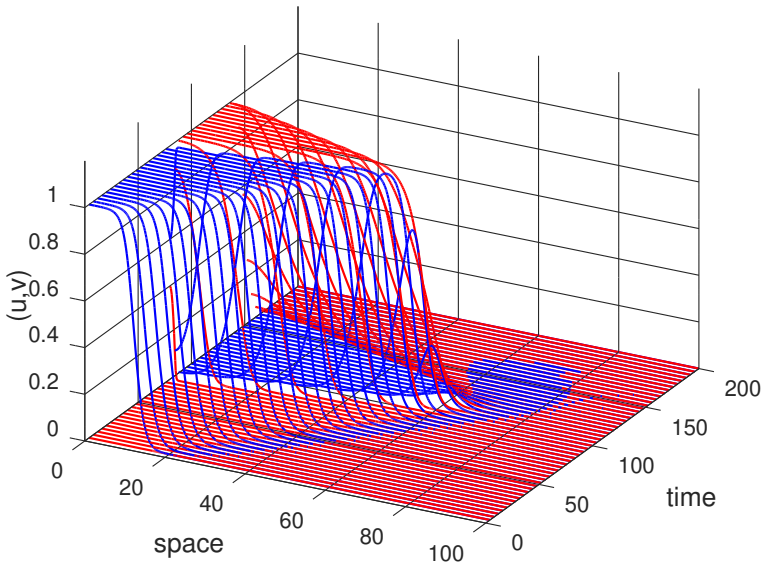
(B) $t = 100$



(C) $t = 120$

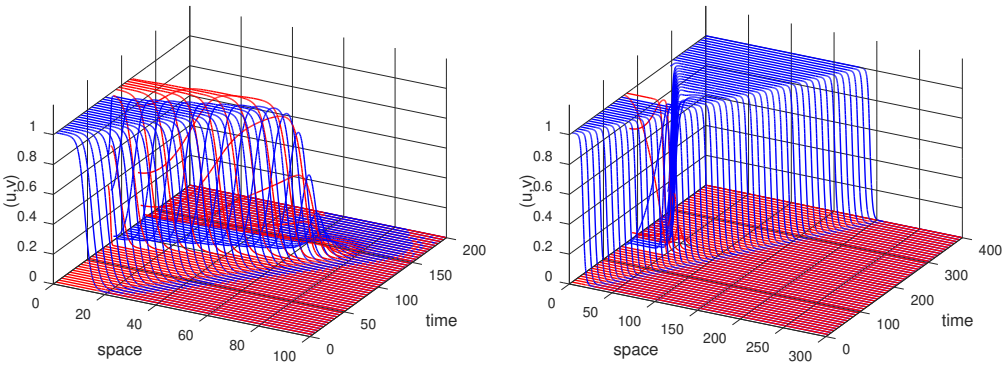


(D) $t = 140$



(E) Space-time representation of the solution (u, v)

FIGURE 1.3. Same as in Figure 1.2, but with $a = 0.55$ and $b = 0.45$.



(A) $a = 0.55$, $b = 0.65$ (B) $a = 0.55$, $b = 0.8$

FIGURE 1.4. Further numerical investigation for a relatively large disadvantage of the brake $b > a > 1/2$. Uniform coextinction remains true for values of b larger than a, but no too large. Joint persistence of both species associated with complicated spatio-temporal behaviours were observed for intermediate values of b ($a = 0.55, b = 0.75$, result not shown).

Next, straighforward algebra shows that any critical point $(u^\star, v^\star) \in \mathbb{R}^2$ satisfies

$$\begin{pmatrix} a & a-(1-h)\,b \\ a-(1-h)\,b & (2h-1)\,b \end{pmatrix} \begin{pmatrix} u^\star \\ v^\star \end{pmatrix} = \begin{pmatrix} a \\ hb \end{pmatrix}.$$

The determinant of the 2×2 matrix above is $-a\,(a-b)-(1-h)^2\,b^2$, which vanishes if and only if $a = b$ and $h = 1$, but in such a case the system reduces to $u^\star + v^\star = 1$ so that there are no critical points in $\mathrm{int}\,(\mathsf{T})$. If $a > b$ or $h < 1$, the 2×2 matrix is invertible and Cramer's rule yields

$$\begin{aligned}(u^\star, v^\star) &= \frac{1}{-a\,(a-b)-(1-h)^2\,b^2} \begin{pmatrix} a(2h-1)b - hb\,(a-(1-h)b) \\ ahb - (a-(1-h)b)\,a \end{pmatrix} \\ &= \frac{1}{a\,(a-b)+(1-h)^2\,b^2} \begin{pmatrix} b(1-h)(a-hb) \\ a(a-b) \end{pmatrix}\end{aligned}$$

This point is in $\mathrm{int}\,(\mathsf{T})$ if and only if $b(1-h)(a-hb) < (1-h)^2b^2$, that is if and only if $a - hb < (1-h)b$, that is if and only if $a < b$. Hence it is not in $\mathrm{int}\,(\mathsf{T})$.

Therefore the minimum of w in T is attained only on the boundary and is $1 - a > 0$. □

Proposition 2.2. *Any solution of (1.1) with initial condition in $\mathscr{C}\left(\mathbb{R}^N, \mathsf{T}\right)$ is valued in T at all times $t \geq 0$.*

Proof. To show that T is an invariant region of the phase space, we use Weinberger's maximum principle [15]. Indeed, the triangle T is a convex invariant set satisfying the so-called slab condition. Therefore we only have to verify that the reaction term is inward-pointing on the boundary of T, namely:

$$\begin{pmatrix} u \\ 0 \end{pmatrix} \circ f\,(u,0) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \geq 0 \text{ for all } u \in [0,1],$$

$$\begin{pmatrix} 0 \\ v \end{pmatrix} \circ f\,(0,v) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \geq 0 \text{ for all } v \in [0,1],$$

$$\begin{pmatrix} u \\ 1-u \end{pmatrix} \circ f\left(u, 1-u\right) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \leq 0 \text{ for all } u \in [0,1].$$

These three conditions are trivially verified (the left-hand side being always zero). □

2.2. **The propagation of the gene drive alone.** In what follows, we fix $v = 0$ and we briefly review some results described in [11] about the dynamics of the gene drive invasion.

If $v = 0$, then (1.1) reduces to

$$\begin{aligned} \mathscr{P}u &= uf_1(u,0) \\ &= u\frac{-au^2 + (3a-1)\,u - (2a-1)}{1-a+a\left(1-u\right)^2} \\ &= \frac{au\left(1-u\right)\left(u - \frac{2a-1}{a}\right)}{1-a+a\left(1-u\right)^2}. \end{aligned}$$

If $a \leq \frac{1}{2}$, this is a monostable equation. Additionally, a bit of algebra shows that this is an equation of KPP type if and only if $a \leq \frac{1}{4}$, the KPP property being here understood in the following weak sense: the maximal growth rate per capita corresponds to sparse populations, namely

$$f_1(0,0) = \max_{u\in[0,1]} f_1(u,0).$$

If $a > \frac{1}{2}$, this is a bistable equation with stable steady states 0 and 1 and unstable intermediate steady state $\theta = \frac{2a-1}{a} \in (0,1)$. All known results on bistable equations, and in particular [6], can therefore be applied to this case.

In particular, the sign of the following quantity plays a crucial role:

$$\int_0^1 \frac{au\left(1-u\right)\left(u-\frac{2a-1}{a}\right)}{1-a+a\left(1-u\right)^2}\mathrm{d}u = \frac{\sqrt{1-a}}{a^{3/2}}\arctan\left(\sqrt{\frac{a}{1-a}}\right) - \frac{1}{2} - \frac{1-a}{a}$$

(the calculation of the integral is not detailed here). Since this is positive if $a = \frac{1}{2}$, negative if $a = 1$ and since

$$\frac{\partial}{\partial a}\left(a \mapsto \arctan\left(\sqrt{\frac{a}{1-a}}\right) - \left(1-\frac{a}{2}\right)\sqrt{\frac{a}{1-a}}\right) = -\frac{(2a-1)\,a^{3/2}}{4a\left(1-a\right)^{3/2}} < 0,$$

there exists a unique $a_0 \in \left(\frac{1}{2}, 1\right)$ such that the integral is positive if $a \in \left[\frac{1}{2}, a_0\right)$, zero if $a = a_0$ and negative if $a \in (a_0, 1]$. It satisfies the numerical approximation $a_0 \simeq 0.6965$.

As a consequence, in the simplified case where $\mathscr{P} = \partial_t - \Delta$ (*i.e.* the total population n is constant), solutions u that are initially compactly supported will always go extinct if $a \in (a_0, 1]$ and will spread and invade if $a \in [0, a_0)$. If $a \in \left(\frac{1}{2}, a_0\right)$, it is also necessary that the initial condition is favorable enough (*i.e.* larger than $\frac{2a-1}{a}$ in a wide region – see the role of the initial data in the emergence of a wave in the bistable case [1], and in particular the existence of bubble-like solutions that can prevent the propagation).

2.3. **Basic phase-plane analysis: spatially uniform stationary states.** Similarly to Subsection 2.2, if $u = 0$, (1.1) reduces to

$$\mathscr{P}v = -bv\frac{(1-v)\left(h(1-v)+v(1-h)\right)}{1-bv^2-2hbv(1-v)}.$$

The right-hand side has exactly the sign of $-v(1-v)$, whence this is a "backward-monostable" equation, where 0 is stable and 1 is unstable.

Again similarly, if $u+v=1$, the equation satisfied by u reduces to

$$\mathscr{P}u = -u\frac{(1-u)\,(1-b+(a-b)\,u)}{1-b-(a-b)u^2}$$

and this is also a "backward-monostable" equation.

All this shows that, regarding the diffusionless system,

- $(0,0)$ is a stable node if $a>\frac{1}{2}$ and is a saddle of $a<\frac{1}{2}$;
- $(1,0)$ is a saddle;
- $(0,1)$ is a saddle.

Moreover,

$$f_2\left(\frac{2a-1}{a},0\right) = \frac{(1-b+a-b)\,(2a-1)+(1-hb)\,(1-a)}{a\left(1-a-a\left(1-\frac{2a-1}{a}\right)^2\right)}.$$

Since this quantity is obviously positive if $a>\frac{1}{2}$, in such a case $(\frac{2a-1}{a},0)$ is an unstable node. Of course, when $a<\frac{1}{2}$, $(\frac{2a-1}{a},0)\notin\mathsf{T}$.

This implies that, for the trajectories we have in mind (*i.e.* contained in the interior of T), convergence to $(1,0)$, $(0,1)$ or $(\frac{2a-1}{a},0)$ is impossible and convergence to $(0,0)$ is possible if and only if $a\geq\frac{1}{2}$. Consequently, Theorem 1.1 is proved.

Any stationary state in the interior of T is a solution of the following algebraic system:

$$\begin{cases} g_1(u,v) = g_2(u,v), \\ g_1(u,v) = w(u,v), \end{cases}$$

which can be solved explicitly (g_1-g_2 is a first-order polynomial while $w-g_1$ is a second-order polynomial). We do not perform this resolution here, as it is tedious and useless. We simply point out that:

- on one hand, the proof of Theorem 1.2 will imply directly the nonexistence of such a coexistence state when $a>\frac{1}{2}$ and $b<b^\star$;
- on the other hand, when $a<\frac{1}{2}$, the flow is rotating anticlockwise on the boundary of T, whence by classical phase-plane arguments (*e.g.* the Poincaré–Bendixson theorem) there exists such a coexistence state. Numerically, we observe that this stationary state is unique and is a stable or unstable spiral (see Figure 2.1).

2.4. **Brake with small selective disadvantage: predator–prey regime.** In what follows it is convenient to understand b as a parameter and to add a subscript b to every object that depends on it. In particular, we rewrite the term $f_b\,(u,v)$ as

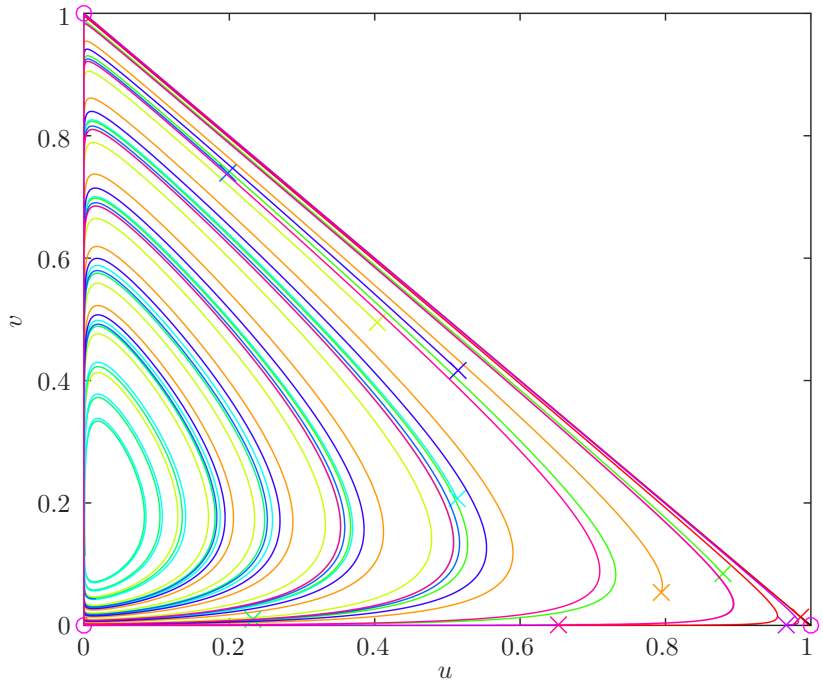
(2.1)

$$f_b\,(u,v) = \frac{w_0\,(u,v)}{w_b\,(u,v)}\left[f_0\,(u,v)+\frac{b}{w_0\,(u,v)}\begin{pmatrix} v\,(v+2h\,(1-u-v)) \\ v\,(v+2h\,(1-u-v))-(v+2u+h\,(1-u-v))\end{pmatrix}\right]$$

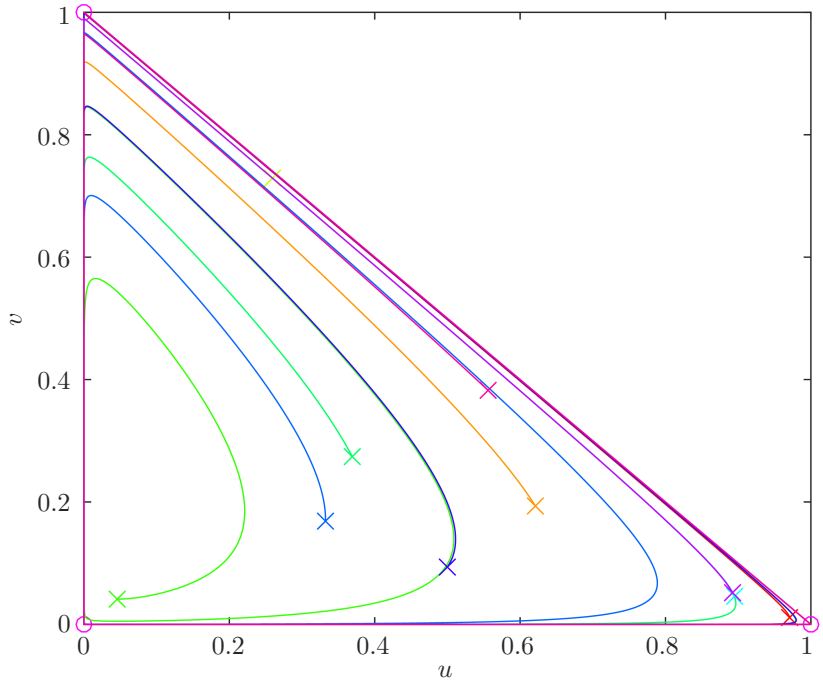
with

$$\begin{aligned} w_0\,(u,v) &= 1-au^2-2au\,(1-u-v) \\ &= 1+au^2-2au+2auv, \end{aligned}$$

$$\begin{aligned} g_0\,(u,v) &= \begin{pmatrix} (1-a)\,u+2\,(1-a)\,(1-u-v) \\ v+2u+(1-u-v)\end{pmatrix} \\ &= \begin{pmatrix} 2\,(1-a)-(1-a)\,u-2\,(1-a)\,v \\ 1+u\end{pmatrix}, \end{aligned}$$



(A) $a = 0.4$, $b = 0.1$, $h = 0.2$.



(B) $a = 0.4$, $b = 0.1$, $h = 0.8$.

FIGURE 2.1. Oscillating trajectories of the homogeneous system with ten random initial conditions and $a < \frac{1}{2}$. Depending on the value of h, the oscillations are either nonperiodic and sustained or damped. For intermediate values of h (not illustrated here), we find a mixture of nonperiodic sustained, periodic sustained and damped oscillations, depending on the initial condition. In all cases, the solutions spend a very long time in the neighborhood of $u = 0$ (final time of the simulation $T = 1000$).

$$\begin{aligned} f_0(u,v) &= \frac{1}{w_0(u,v)} g_0(u,v) - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{w_0(u,v)} \begin{pmatrix} -(2a-1) + (3a-1)u - au^2 - 2v(1-a+au) \\ u(1+2a-au-2av) \end{pmatrix} \\ &= \frac{1}{w_0(u,v)} \begin{pmatrix} a(1-u)(u-\theta) - 2v(1-a+au) \\ u(1+2a-au-2av) \end{pmatrix}. \end{aligned}$$

Proposition 2.3. *There exists $\overline{b}_1 \in (0,1]$ such that, if $b \in \left[0, \overline{b}_1\right]$, then (1.1) has a predator–prey structure where u is the prey and v is the predator, namely*

$$\partial_v f_{b,1} < 0 \text{ and } \partial_u f_{b,2} > 0 \text{ in } \mathsf{T}.$$

Proof. First, we study the case of $b = 0$.

By straightforward algebra and definition of θ,

$$\begin{aligned} \partial_v f_{0,1}(u,v) &= \frac{2}{w_0(u,v)^2} \left(-(1-a+au) w_0(u,v) - (a(1-u)(u-\theta) - 2v(1-a+au)) au\right) \\ &= \frac{2}{w_0(u,v)^2} \left(-2a(2a-1)u^2 - (1-a)\right) \\ &\leq -\frac{2}{w_0(u,v)^2} \min_{z\in[0,1]} \left(2a(2a-1)z^2 + (1-a)\right). \end{aligned}$$

On one hand, if $a \geq \frac{1}{2}$, the minimum above is $1-a$. On the other hand, if $a < \frac{1}{2}$, the minimum above is $-2a(1-2a) + 1 - a = 4a^2 - 3a + 1 > 0$. In all cases, since $w_0 \leq 1$ in T,

$$\partial_v f_{0,1}(u,v) \leq -\frac{2}{w_0(u,v)^2} \min\left(1-a, 4a^2-3a+1\right) \leq -2\min\left(1-a, 4a^2-3a+1\right) < 0.$$

Similarly,

$$\begin{aligned} \partial_u f_{0,2}(u,v) &= \frac{w_0(u,v)(1+2a(1-u-v)) + 2au(1+au+2a(1-u-v))(1-u-v)}{w_0(u,v)^2} \\ &\geq \frac{1-a}{w_0(u,v)^2} \geq 1-a > 0. \end{aligned}$$

Thus (1.1) in the limiting case $b = 0$ has indeed the claimed predator–prey structure.

Finally, thanks to the smooth convergence of f_b to f_0 as $b \to 0$ and the preceding uniform estimates, the existence of $\overline{b}_1 > 0$ as in the statement is immediate. □

3. Proof of Theorem 1.2

In this subsection we assume $a > \frac{1}{2}$ so the drive has a bistable dynamic. We denote $\theta = \frac{2a-1}{a}$ its intermediate steady state.

Interestingly, when $b \leq \overline{b}_1$, the system under consideration has the structure of a predator–prey system with Allee effect on the prey. Such systems have been studied in the literature but, apart from numerical simulations that show an incredibly wide and complicated variety of behaviors, almost nothing is known [7, 10, 14, among others]. Fortunately, the analysis of our particular case is indeed possible, thanks to a very simple phase-plane structure (that can easily be observed numerically).

We will use as in Subsection 2.4 the subscripts b. In order to study more precisely the role of b in f_b, we define additionally

$$\begin{aligned} r(u,v) &= \frac{1}{v}\begin{pmatrix} u \\ v \end{pmatrix} \circ \begin{pmatrix} v\left(v+2h\left(1-u-v\right)\right) \\ v\left(v+2h\left(1-u-v\right)\right)-\left(v+2u+h\left(1-u-v\right)\right) \end{pmatrix} \\ &= \begin{pmatrix} u\left(v+2h\left(1-u-v\right)\right) \\ v\left(v+2h\left(1-u-v\right)\right)-\left(v+2u+h\left(1-u-v\right)\right) \end{pmatrix} \\ &= \begin{pmatrix} u\left(2h-2hu+\left(1-2h\right)v\right) \\ \left(1-2h\right)v^2+\left(3h-1\right)v-h-2huv-\left(2-h\right)u \end{pmatrix} \\ &= \begin{pmatrix} u\left(v+2h\left(1-u-v\right)\right) \\ \left(v-1\right)\left(\left(1-2h\right)v+h\right)-u\left(2hv+2-h\right) \end{pmatrix}. \end{aligned}$$

Thus the reaction term of (1.1) reads

$$\begin{pmatrix} u \\ v \end{pmatrix} \circ f_b\left(u,v\right) = \frac{w_0\left(u,v\right)}{w_b\left(u,v\right)}\left[\begin{pmatrix} u \\ v \end{pmatrix} \circ f_0\left(u,v\right) + \frac{bv}{w_0\left(u,v\right)} r\left(u,v\right)\right]. \tag{3.1}$$

Remark. The function r satisfies, in T, $r_1 \geq 0$ with equality if and only if $u = 0$ or $(u,v) = (1,0)$ and $r_2 \leq 0$ with equality if and only if $(u,v) = (0,1)$. This is mostly obvious but we point out that $(1-2h)v + h = h(1-v) + v(1-h) \geq 0$ cannot vanish if $v \in (0,1)$.

3.1. **Three geometric lemmas.**

Lemma 3.1. *There exists $\bar{b}_2 \in (0,1]$ such that, if $b \in \left[0, \bar{b}_2\right]$, then for all $\mu \geq 0$ and all $u \in [\theta, 1]$ such that $\mu\left(u-\theta\right) \leq 1-u$,*

$$\left(\begin{pmatrix} u \\ \mu\left(u-\theta\right) \end{pmatrix} \circ f_b\left(u, \mu\left(u-\theta\right)\right)\right) \cdot \begin{pmatrix} \mu \\ -1 \end{pmatrix} \leq 0. \tag{3.2}$$

Proof. Let

$$\mathsf{C} = \left\{(\mu,u) \in [0,+\infty) \times [\theta,1] \;\middle|\; \mu\left(u-\theta\right) \leq 1-u\right\}.$$

In view of (3.1), (3.2) is equivalent to

$$\left(\begin{pmatrix} u \\ u-\theta \end{pmatrix} \circ f_0\left(u, \mu\left(u-\theta\right)\right) + \frac{b\left(u-\theta\right)}{w_0\left(u, \mu\left(u-\theta\right)\right)}\begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix}\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \leq 0.$$

Straightforward algebra and the definition of θ lead to

$$\begin{pmatrix} u \\ u-\theta \end{pmatrix} \circ f_0\left(u, \mu\left(u-\theta\right)\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -\frac{u\left(u-\theta\right)}{w_0\left(u, \mu\left(u-\theta\right)\right)}\left(1+a+2a\mu\right)$$

whence (3.2) is actually equivalent to

$$-u\left(1+a+2a\mu\right) + b\begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \leq 0.$$

Clearly, this is true for all $\mu \geq 0$ and all $u \in [\theta, 1]$ if $b = 0$. Hence from now on we focus on the case $b > 0$. In such a case, requiring (3.2) for all $(\mu, u) \in \mathsf{C}$ is equivalent to requiring

$$\frac{1}{b} \geq \sup_{(\mu,u)\in\mathsf{C}} \frac{\max\left(0, \begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)}{u\left(1+a+2a\mu\right)}.$$

In view of the signs of r_1 and r_2 in T, the right-hand side is positive, so that the above inequality reduces to

$$\frac{1}{b} \geq \sup_{(\mu,u)\in\mathsf{C}} \frac{\begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}}{u\left(1+a+2a\mu\right)} > 0.$$

It only remains to verify the finiteness of this supremum.

Let $(\mu, u) \in \mathsf{C}$. From

$$\mu r_1\left(u, \mu\left(u-\theta\right)\right) = \left(\mu - 2h\left(\mu+1\right)\right)\mu u^2 + \left(2h + 2h\mu\theta - \mu\theta\right)\mu u$$

and

$$\begin{aligned} -r_2\left(u, \mu\left(u-\theta\right)\right) = &-\left(\mu - 2h\left(\mu+1\right)\right)\mu u^2 - \left(\mu\left(4h\mu\theta + 3h + 2h\theta - 1 - 2\mu\theta\right) + h - 2\right)u \\ &- \left(\mu^2\theta^2 + \mu\theta - 2h\mu^2\theta^2 - 3h\mu\theta - h\right), \end{aligned}$$

it follows

$$\begin{aligned} \begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = &\left((1-2h)\mu^2\theta - (h + 2h\theta - 1)\mu + 2 - h\right)u \\ &- \left((1-2h)\mu^2\theta^2 + (1-3h)\mu\theta - h\right), \end{aligned}$$

which reads, as a polynomial of μ,

$$\begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (1-2h)\,\theta\,(u-\theta)\,\mu^2 + ((3h-1)\,\theta - (2h\theta + h - 1)\,u)\,\mu + (2-h)\,u + h.$$

On one hand, if $h > \frac{1}{2}$, the above second-order polynomial of μ is bounded above by some constant, whence

$$\sup_{(\mu,u)\in\mathsf{C}} \frac{\begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}}{u\left(1+a+2a\mu\right)} < +\infty.$$

On the other hand, if $h \leq \frac{1}{2}$, then we use $\mu\left(u-\theta\right) \leq 1-u$ and obtain

$$\begin{aligned} \begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} &\leq (1-2h)\,\theta\mu\,(1-u) + ((3h-1)\,\theta - (2h\theta + h - 1)\,u)\,\mu + (2-h)\,u + h \\ &\leq \left((1-\theta)\,u - h\,(u-\theta)\right)\mu + (2-h)\,u + h, \end{aligned}$$

whence

$$\sup_{(\mu,u)\in\mathsf{C}} \frac{\begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}}{u\left(1+a+2a\mu\right)} < +\infty$$

is again true.

The proof is ended with

$$\bar{b}_2 = \min\left(1, \left(\sup_{(\mu,u)\in\mathsf{C}} \frac{\begin{pmatrix} \mu r_1\left(u, \mu\left(u-\theta\right)\right) \\ r_2\left(u, \mu\left(u-\theta\right)\right) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}}{u\left(1+a+2a\mu\right)}\right)^{-1}\right).$$

□

By very similar considerations, we can also prove the following lemma. For the sake of brevity, we do not detail the proof.

Lemma 3.2. *There exists $\overline{b}_3 \in (0,1]$ such that, if $b \in \left[0, \overline{b}_3\right]$, then for all $v \in [0, 1-\theta]$,*

$$\left(\begin{pmatrix}\theta\\ v\end{pmatrix} \circ f_b\left(\theta, v\right)\right) \cdot \begin{pmatrix}1\\ 0\end{pmatrix} \leq 0.$$

Next, using the fact that $(\theta, 0)$ is an unstable node (see Subsection 2.3), we can prove the following similar lemma.

Lemma 3.3. *There exists $\overline{\eta} \in \left(0, \frac{1-\theta}{4}\right)$ such that, for all $\eta \in (0, \overline{\eta})$ and all $\mu > 0$:*

(1) for all $u \in [\theta - \eta, \theta + \eta]$,

$$\left(\begin{pmatrix}u\\ \min\left(1, \frac{\mu}{2}\right)(u - \theta + \eta)\end{pmatrix} \circ f_b\left(u, \min\left(1, \frac{\mu}{2}\right)(u - \theta + \eta)\right)\right) \cdot \begin{pmatrix}\min\left(1, \frac{\mu}{2}\right)\\ -1\end{pmatrix} \leq 0.$$

(2) the straight lines of equation $v = \mu\left(u - \theta\right)$ and $v = \min\left(1, \frac{\mu}{2}\right)(u - \theta + \eta)$ intersect at $u = \min\left(\theta + \eta, \theta + \frac{\eta}{\mu - 1}\right)$ with

$$\left(\min\left(\theta + \eta, \theta + \frac{\eta}{\mu - 1}\right), \mu\left(\min\left(\theta + \eta, \theta + \frac{\eta}{\mu - 1}\right) - \theta\right)\right) \in \operatorname{int}\left(\mathsf{T}\right).$$

These lemmas will be used to construct a family of convex sets, illustrated in Figure 3.1.

3.2. **Main proof.** We are now in position to prove the main theorem.

Theorem 3.4. *Assume $b \leq \min\left(\overline{b}_1, \overline{b}_2, \overline{b}_3\right)$.*

Let (u, v) be the solution of (1.1) with initial condition $(u_0, v_0) \in \mathscr{C}\left(\mathbb{R}^N, \mathsf{T}\right)$ satisfying

$$v_0 \neq 0 \quad \textit{and} \quad \lim_{\|x\| \to +\infty} \left(u_0, v_0\right)(x) = (0, 0).$$

Then co-extinction occurs, namely

$$\lim_{t \to +\infty} \left(\sup_{x \in \mathbb{R}^N} u\left(t, x\right) + \sup_{x \in \mathbb{R}^N} v\left(t, x\right)\right) = 0.$$

Proof. First, for all $\mu \geq 0$, we define the line segment

$$\mathsf{S}_\mu = \left\{(\tilde{u}, \tilde{v}) \in \mathsf{T} \mid \tilde{v} = \mu\left(\tilde{u} - \theta\right)\right\}.$$

For all $t \geq 0$, we define the curve

$$\mathsf{G}_t = \left\{(u, v)\left(t, x\right) \mid x \in \mathbb{R}^N\right\} \subset \mathsf{T}.$$

For all $t \geq 0$ and $\mu \geq 0$, we define the (Euclidean) distance between G_t and S_μ

$$d\left(t, \mu\right) = \operatorname{dist}\left(\mathsf{G}_t, \mathsf{S}_\mu\right).$$

By standard parabolic estimates and the fact that $(0, 0)$ is a solution of (1.1), the assumptions imply

$$\lim_{\|x\| \to \infty} \left(u, v\right)\left(t, x\right) = (0, 0) \text{ for all } t \geq 0.$$

Consequently, for all $t \geq 0$ and $\mu \geq 0$, there exists $x_{t,\mu} \in \mathbb{R}^N$ such that

$$d(t, \mu) = \operatorname{dist}\left(\left(u, v\right)\left(t, x_{t,\mu}\right), \mathsf{S}_\mu\right).$$

Now, let

$$T = \sup\left\{t \geq 0 \mid \mathsf{G}_t \cap (\theta, 1] \times [0, 1] \neq \emptyset\right\}$$

and let us show that $T < +\infty$. Without loss of generality, we restrict the analysis to the case $T > 0$.

For all $t \in [0, T]$, we can define

$$\mu_t = \inf\left\{\mu \geq 0 \mid d\left(t, \mu\right) = 0\right\}$$

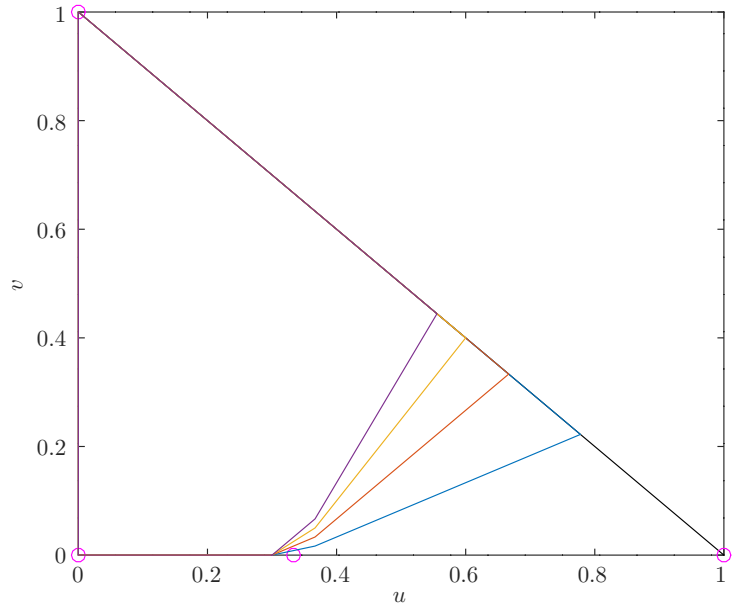


FIGURE 3.1. Examples of sets C_μ with $\mu \in \{0.5, 1, 1.5, 2\}$. (The value $\eta = \frac{\theta}{10}$ was chosen here for graphic clarity and was not numerically checked.)

and, by continuity, we find

$$d(t, \mu_t) = \operatorname{dist}\left((u, v)(t, x_{t,\mu_t}), \mathsf{S}_{\mu_t}\right) = 0.$$

Next, by virtue of Weinberger's maximum principle [15] applied in the convex invariant set T satisfying the so-called slab condition, for all $t \in (0, T]$, $(u, v)(t, x_{t,\mu_t}) \in \operatorname{int}(\mathsf{T})$. This directly implies that for all $t \in (0, T]$, $\mu_t > 0$.

However, in view of Lemma 3.1 and Lemma 3.3, for any $\eta \in (0, \overline{\eta})$ and any $\mu > 0$, the convex set satisfying a slab condition

$$\mathsf{C}_\mu = \left\{(\tilde{u}, \tilde{v}) \in \mathsf{T} \mid \tilde{v} \geq \mu(\tilde{u} - \theta), \tilde{v} \geq \min\left(1, \frac{\mu}{2}\right)(\tilde{u} - \theta + \eta)\right\}$$

is again an invariant set for (1.1). Now we define $T_0 = \min\left(1, \frac{T}{2}\right)$ and fix $\eta \in (0, \overline{\eta})$ so small that

$$\mathsf{G}_{T_0} \subset \mathsf{C}_{\mu_{T_0}}.$$

Applying again Weinberger's maximum principle, we find that for all $(t, t') \in [T_0, T]^2$ such that $t' > t > T_0$, $(u, v)\left(t', x_{t',\mu_{t'}}\right) \in \operatorname{int}(\mathsf{C}_{\mu_t})$. Hence the family $(\mathsf{C}_{\mu_t})_{t \geq T_0}$ is decreasing with respect to the inclusion, or in other words the family $(\mu_t)_{t \geq T_0}$ is increasing.

Assuming by contradiction that $T = +\infty$, we can define

$$\mu_\infty = \lim_{t \to +\infty} \mu_t \in (\mu_{T_0}, +\infty].$$

Then, by standard parabolic estimates, the family

$$\left((t, x) \mapsto (u, v)(t + n, x + x_{n,\mu_n})\right)_{n \in \mathbb{N}}$$

converges locally uniformly up to extraction to an entire solution (u_∞, v_∞) of (1.1) valued in C_{μ_∞} and satisfying $(u_\infty, v_\infty)(0,0) \in \mathsf{S}_{\mu_\infty}$, where, obviously,

$$\mathsf{C}_\infty = \{(\tilde{u}, \tilde{v}) \in \mathsf{T} \mid \tilde{u} \leq \theta, \tilde{v} \geq (\tilde{u} - \theta + \eta)\}$$

and

$$\mathsf{S}_\infty = \{\theta\} \times [0, 1-\theta].$$

Applying again Weinberger's maximum principle and using the fact that $(\theta, 0) \notin \mathsf{C}_{\mu_\infty}$, we find that necessarily $\mu_\infty = +\infty$. However, by the scalar comparison principle and the predator-prey structure (see Proposition 2.3), u_∞ is a subsolution for the bistable equation

$$\mathscr{P}\tilde{u} = \tilde{u} f_{b,1}(\tilde{u}, 0)$$

satisfying $u_\infty \leq \theta$ with $u_\infty(0,0) = \theta$. This directly implies $u_\infty = \theta$, and then back to the system $v_\infty = 0$, which is a contradiction.

Therefore $T < +\infty$, that is (u, v) enters in finite time the domain C_∞. Again by Weinberger's maximum principle applied this time to the convex set given by Lemma 3.2, it actually enters in finite time the interior of C_∞. Repeating the comparison with the scalar bistable equation

$$\mathscr{P}\tilde{u} = \tilde{u} f_{b,1}(\tilde{u}, 0),$$

it follows that $u \to 0$ uniformly in space asymptotically in time.

Quite interestingly, the convergence of v to 0 does not follow from classical predator–prey considerations. Indeed, in our case, $(0,1)$ is a steady state, and moreover it is a saddle that attracts the solutions $(\tilde{u}, \tilde{v})$ such that $\tilde{u} + \tilde{v} = 1$. Fortunately, (u, v) does not belong to this special "wild-type free" class of solutions and, with some additional care, we will be able to conclude as expected.

In view of the convergence of u, the function

$$\varepsilon : t \mapsto \sup_{t' \geq t} \left(\max_{x \in \mathbb{R}^N} u(t', x) \right)$$

converges monotonically to 0. Thanks to the predator–prey structure, it follows just as before that, for all $t_0 \geq 0$, v is in $[t_0, +\infty) \times \mathbb{R}^N$ a subsolution for the equation

$$\mathscr{P}\tilde{v} = \tilde{v} f_{b,2}(\varepsilon(t_0), \tilde{v}).$$

The Taylor expansion of $f_{b,2}$ and the predator–prey structure bring forth the existence of a positive constant $C > 0$ such that

$$\begin{aligned}
\mathscr{P}v &\leq v \left[\frac{b(-h + (2h-1)v)(1-v)}{1 - 2hbv + (2h-1)bv^2} + \partial_u f_{b,2}(0, v)\varepsilon(t_0) + o(\varepsilon(t_0)) \right] \\
&\leq v \left[\frac{b(-h + (2h-1)v)(1-v)}{1 - 2hbv + (2h-1)bv^2} + C\varepsilon(t_0) \right].
\end{aligned}$$

Now, thanks to $w_b \leq 1$ and $-h + (2h-1)v \leq 0$,

$$\frac{b(-h + (2h-1)v)(1-v)}{1 - 2hbv + (2h-1)bv^2} \leq -b(h - (2h-1)v)(1-v).$$

If $h = 1$, we find

$$\frac{b(-h + (2h-1)v)(1-v)}{1 - 2hbv + (2h-1)bv^2} \leq -b(1-v)^2.$$

If $h \in \left(\frac{1}{2}, 1\right)$, we find

$$\frac{b(-h + (2h-1)v)(1-v)}{1 - 2hbv + (2h-1)bv^2} \leq -b(1-h)(1-v).$$

If $h \in \left(0, \frac{1}{2}\right)$, we find

$$\frac{b\left(-h+(2h-1)\,v\right)(1-v)}{1-2hbv+(2h-1)\,bv^2} \leq -bh\,(1-v)\,.$$

If $h = 0$, we find

$$\frac{b\left(-h+(2h-1)\,v\right)(1-v)}{1-2hbv+(2h-1)\,bv^2} \leq -bv\,(1-v)\,.$$

Hence the function $\hat{v} = 1 - v$ satisfies

$$\mathscr{P}\hat{v} \geq (1-\hat{v})\left(F\left(\hat{v}\right) - C\varepsilon\left(t_0\right)\right),$$

where

$$F : z \in [0,1] \mapsto \begin{cases} bz^2 & \text{if } h = 1, \\ b\,(1-h)\,z & \text{if } h \in \left(\frac{1}{2}, 1\right) \\ bhz & \text{if } h \in \left(0, \frac{1}{2}\right) \\ bz\,(1-z) & \text{if } h = 0. \end{cases}$$

In both cases, we can Lipschitz-continuously extend F in $[-1,1]$ as an even nonnegative nonzero function which is increasing in $\left(0, \frac{1}{2}\right)$ and maps $\left[0, \frac{1}{2}\right]$ onto $\left[0, F\left(\frac{1}{2}\right)\right]$. Assume that t_0 is so large that $C\varepsilon\left(t_0\right) < F\left(\frac{1}{2}\right)$ and let $z_0 \in \left(0, \frac{1}{2}\right)$ be the root of $F(z) = C\varepsilon(t_0) > 0$.

To conclude, we distinguish two cases, depending on whether $h = 0$ or not.

On one hand, if $h > 0$, the reaction term $(1-\hat{v})\left(F\left(\hat{v}\right) - C\varepsilon\left(t_0\right)\right)$ has in $[-z_0, 1]$ a bistable structure: its only zeros are $-z_0$, $+z_0$ and 1 with negative derivative at $-z_0$ and 1 and positive derivative at $+z_0$. By taking t_0 large enough, we can further assume that

$$\int_{-z_0}^{1} (1-z)\left(F(z) - C\varepsilon\left(t_0\right)\right) \mathrm{d}z > 0.$$

By well-known results on bistable equations, this condition implies that the solution of

$$\begin{cases} \mathscr{P}z = (1-z)\left(F\left(z\right) - C\varepsilon\left(t_0\right)\right) \\ z(t_0, x) = \hat{v}\left(t_0, x\right) \text{ for all } x \in \mathbb{R} \end{cases}$$

converges uniformly in space to 1 as $t \to +\infty$. Here,

$$\lim_{\|x\| \to \infty} v\left(t_0, x\right) = 0$$

(established at the beginning of the proof) is crucial. Finally, by the comparison principle, $\hat{v}$ converges uniformly to 1, or in other words v converges uniformly to 0.

On the other hand, if $h = 0$, then $(1-\hat{v})\left(F\left(\hat{v}\right) - C\varepsilon\left(t_0\right)\right)$ has in $[-z_0, 1]$ one more zero, say z_1, which is in $(z_0, 1)$. The steady states z_0 and 1 are unstable whereas 0 and z_1 are stable. By arguments similar to the previous case, we obtain

$$\lim_{t \to +\infty} \inf_{x \in \mathbb{R}^N} \hat{v}(t,x) \geq z_1.$$

But z_1 depends continuously on t_0 and converges to 1 as $t_0 \to +\infty$. Passing to the limit, we deduce

$$\lim_{t \to +\infty} \inf_{x \in \mathbb{R}^N} \hat{v}(t,x) \geq 1,$$

whence v converges uniformly to 0 indeed. □

4. Discussion

4.1. **On different choices of selective disadvantages.** In what follows, we give some evidence arguing in favor of co-extinction when $a > \frac{1}{2}$ (and $b \leq a$), versus the existence of a joint front with complex spatio-temporal dynamics when $a < \frac{1}{2}$, see also Figure 1.4.

4.1.1. *Monostable drive and brake with small selective disadvantage.* In this subsection we assume $a \leq \frac{1}{2}$ so that the drive has a monostable dynamic and $b \leq \overline{b}_1$ so that we are in the predator–prey regime (see Proposition 2.3). We recall that, by Subsection 2.3, global convergence to $(0,0)$ for the diffusionless system cannot occur if $a < \frac{1}{2}$, as it is a saddle. As showed previously by Figure 2.1, the system is oscillating with either damped oscillations and eventually convergence to a coexistence state or sustained oscillations, possibly approaching the heteroclinic orbits along the boundary. What about the reaction–diffusion system? In particular, is the monostable drive fast enough to persist ahead of the oscillating front?

First, since $v \mapsto f_2(1,v)$ is

$$v \mapsto \frac{2(1-b) - b(1-h)v}{1-a+(2h-1)bv^2 + 2(a-b)v} - 1$$

and is therefore decreasing with respect to v, the nonlinearity $vf_2(1,v)$ has a KPP structure.

Hence (1.1) is a particular case of the more general framework investigated by Ducrot–Giletti–Matano [3], provided that $a \leq \frac{1}{4}$ (the prey follows KPP-like dynamics), and the drift term due to a heterogeneous population size is neglected, *i.e.* $\mathscr{P} = \partial_t - \Delta$. However, regarding the part of their analysis we are interested in here, we believe the KPP condition and the absence of drift could be relaxed. In any case, the following discussion can be rigorously justified when $a \leq \frac{1}{4}$ and n is constant following the arguments in [3].

This boils down to the comparison between the spreading speed c of u invading the wild-type type 0 alone (cf. Subsection 2.2) and the spreading speed of v invading the gene drive $u \sim 1$ separately, namely $2\sqrt{f_2(1,0)} = 2\sqrt{\frac{1+a-2b}{1-a}}$.

In the case $a \leq \frac{1}{4}$, the sole gene drive verifies a KPP condition ensuring that the propagation occurs at the explicit speed $2\sqrt{f_1(0,0)} = 2\sqrt{1-2a}$. It is immediate to see that $2\sqrt{1-2a} < 2\sqrt{\frac{1+a-2b}{1-a}}$, so that u is unable to evade, and will eventually be caught up by v.

In the case $a \in \left(\frac{1}{4}, \frac{1}{2}\right)$, the spreading speed c can be estimated by comparison with a well-chosen KPP equation, namely with a KPP equation whose reaction term is larger than or equal to $u \mapsto uf_1(u,0)$ in $[0,1]$. It can be verified easily that the minimal KPP reaction term satisfying this condition is exactly defined in $[0,1]$ as

$$u \mapsto u \max_{z \in [u,1]} f_1(z,0).$$

Denoting

$$\alpha = \max_{u \in [0,1]} f_1(u,0),$$

the KPP speed associated with this reaction term is exactly $2\sqrt{\alpha}$, so that the spreading speed c of u invading 0 satisfies $c \leq 2\sqrt{\alpha}$. To estimate α itself, we use the fact that it is a double root of the second-order polynomial equation

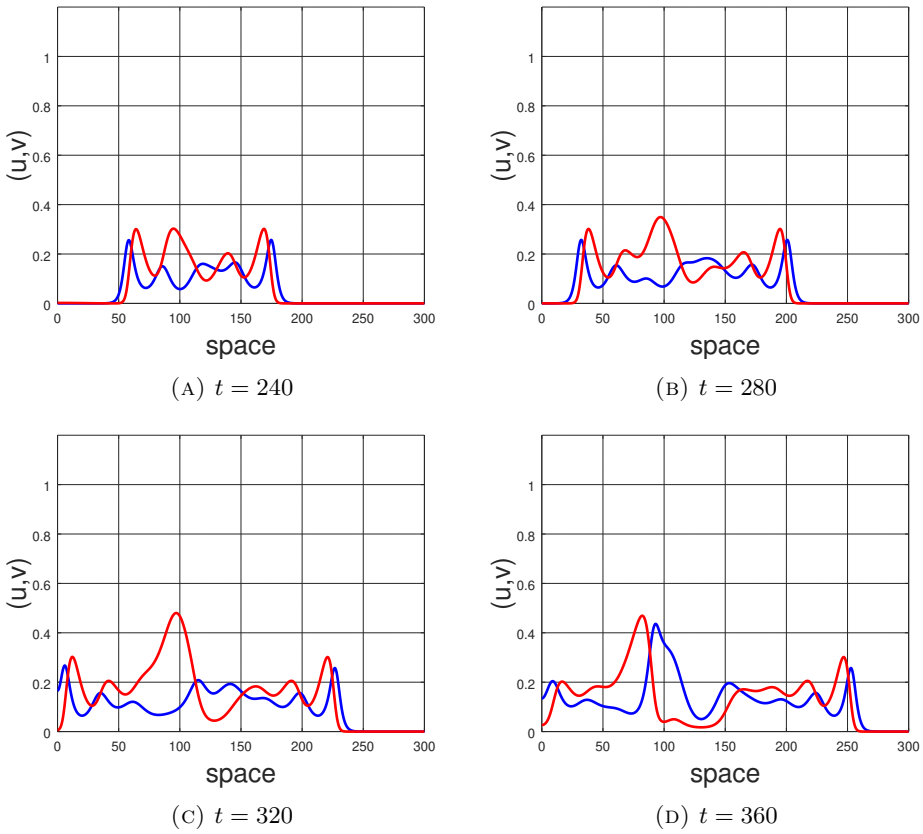
$$\alpha\left(1-a+a\left(1-u\right)^2\right) = a\left(1-u\right)\left(u-\theta\right).$$

After some algebra, we find that the discriminant of this equation vanishes if and only if $\alpha = \frac{1-\sqrt{a}}{2\sqrt{a}}$, whence the estimate $c \leq 2\sqrt{\alpha}$ finally reads

$$c \leq \sqrt{2\frac{1-\sqrt{a}}{\sqrt{a}}}.$$

Subsequently, we consider the quantity

$$2\sqrt{\frac{1+a-2b}{1-a}} - \sqrt{2\frac{1-\sqrt{a}}{\sqrt{a}}}.$$



(A) $t = 240$ (B) $t = 280$

(C) $t = 320$ (D) $t = 360$

FIGURE 4.1. Further numerical snapshots of the solution of (1.1) in the case $a = 0.45$, and $b = 0.35$, $h = 0.5$, and n is spatially homogeneous, see Figure 1.2.

It turns out that there exists $a_{1,b} \in \left(0, \frac{1}{2}\right)$, increasing with respect to $b \in [0, a]$ and satisfying $a_{1,a} = \frac{1}{9}$, such that this quantity is positive if and only if $a > a_{1,b}$. Consequently, if $a \geq \frac{1}{4} > a_{1,b}$, we expect that u is never able to evade v and a joint invasion front is observed.

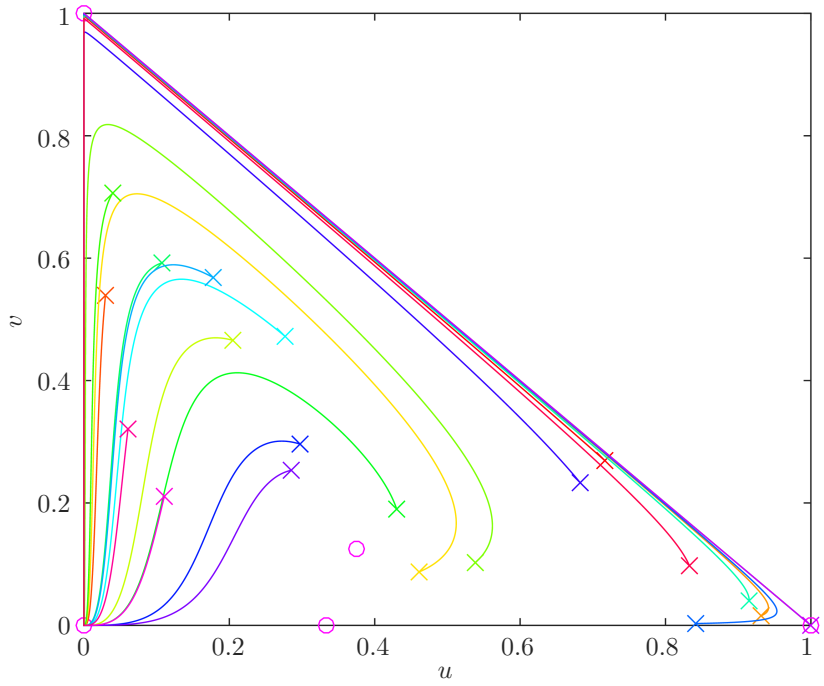
Due to the diffusion and possible Turing instabilities, we cannot really expect simple sustained or damped oscillations in the wake of the joint invasion front. Numerically, we observe very complicated spatio-temporal patterns (see Figure 4.1).

4.1.2. *Brake and drive with the same large selective disadvantage.* When b is close to $a > \frac{1}{2}$, the dynamics are more complicated.

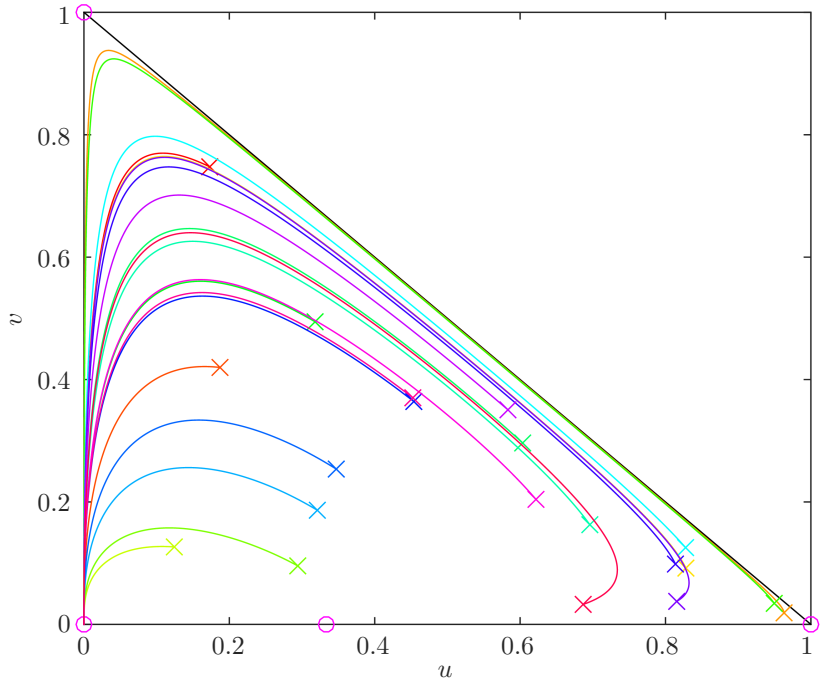
In the limiting case $a = b$ with $h = 1$, the following properties can be established (quite directly). Of course they persist in a neighborhood of $a = b$, $h = 1$.

(1) The dynamics in T are no longer of predator–prey type. The system remains predative outside of the neighborhood of $(0, 0)$ delimited by the graph of

$$u \mapsto \frac{1}{2}\left(2 - u - \sqrt{(2-u)^2 + 4(u-\theta)}\right),$$



(A) $a = 0.6$, $b = 0.6$, $h = 1$.



(B) $a = 0.6$, $b = 0.6$, $h = 0.1$.

FIGURE 4.2. Trajectories of the diffusionless system with twenty random initial conditions and $b = a > \frac{1}{2}$. When $h = 1$, there is a coexistence state that is an unstable spiral. In both cases, $(0, 0)$ seems to be globally attractive.

but inside this neighborhood, the system is now cooperative. Even though $(0,0)$ might be globally attractive inside the cooperative region (and this should actually be easy to establish), general solutions of the reaction–diffusion system might never enter uniformly this region. In other words, it is not sufficient anymore to prove that $u < \theta$ is satisfied in finite time. In order to prove that the solution enters in finite time the cooperative region, we need a Lyapunov function or a family of contractant sets, as in the proof of Theorem 1.2. We did not manage to perform such a construction.

(2) If $a < \frac{2}{3}$, there is a (unique) spatially uniform interior stationary state at

$$(u^\star, v^\star) = \left(\frac{a}{4(1-a)}, \frac{2-3a}{4(1-a)} \right).$$

In view of Figure 4.2, this steady state is, regarding the diffusionless system, an unstable spiral. This is one of the main obstacles encountered when trying to construct a Lyapunov function or a family a contractant sets.

Because of these obstacles, we did not manage to find an analytical proof of the coextinction. Nevertheless, in view of numerical experiments, it remains conjectured to hold true.

4.2. **On imperfect conversion efficiencies.** When the conversion efficiencies of the drive and the brake are no longer assumed to be perfect, the system reads instead

$$\begin{cases} \mathscr{P}u = u \left(\dfrac{(1-a)\,u + (1-c_B)\,(1-h_{BD}b)\,v + (2c_D\,(1-a) + (1-c_D)\,(1-h_{DO}a))\,(1-u-v)}{w_{c_D,c_B}\,(u,v)} - 1 \right), \\ \mathscr{P}v = v \left(\dfrac{(1-b)\,v + (2c_B\,(1-b) + (1-c_B)\,(1-h_{BD}b))\,u + (1-hb)\,(1-u-v)}{w_{c_D,c_B}\,(u,v)} - 1 \right). \end{cases}$$

Here c_D, c_B, h_{DO} and h_{BD} are all in $[0,1]$ and are respectively the conversion efficiency of the drive, the conversion efficiency of the brake, the dominance of D on O and the dominance of B on D. The normalizing mean fitness $w_{c_D,c_B}(u,v)$ is, similarly to the numerator, a mere algebraic modification of the perfect case.

It is therefore quite clear that the various uniform estimates on the reaction term we derived near the limit $b \sim 0$ should remain true for values of (c_D, c_B) close enough to $(1,1)$ and up to a slight modification of the threshold $a = \frac{1}{2}$. Hence both Theorem 1.1 and Theorem 1.2 are expected to remain true in this framework. We leave this extension for future work.

References

[1] Nick H. Barton and Michael Turelli. Spatial waves of advance with bistable dynamics: Cytoplasmic and genetic analogues of Allee effects. *The American Naturalist*, 178(3):–48, 2011. PMID: 21828986.

[2] Andrea Beaghton, Pantelis John Beaghton, and Austin Burt. Gene drive through a landscape: Reaction–diffusion models of population suppression and elimination by a sex ratio distorter. *Theoretical Population Biology*, 108:51–69, 2016.

[3] Arnaud Ducrot, Thomas Giletti, and Hiroshi Matano. Spreading speeds for multidimensional reaction-diffusion systems of the prey-predator type. *in preparation.*

[4] Kevin M. Esvelt and Neil J. Gemmell. Conservation demands safe gene drive. *PLOS Biology*, 15(11):1–8, 11 2017.

[5] Kevin M. Esvelt, Andrea L. Smidler, Flaminia Catteruccia, and George M. Church. Emerging Technology: Concerning RNA-guided gene drives for the alteration of wild populations. *eLife*, 3:e03401, jul 2014.

[6] Paul C. Fife and J. B. McLeod. The approach of solutions of nonlinear diffusion equations to travelling front solutions. *Archive for Rational Mechanics and Analysis*, 65(4):335–361, 1977.

[7] Andrew Morozov, Sergei Petrovskii, and Bai–Lian Li. Bifurcations and chaos in a predator-prey system with the allee effect. *Proceedings of the Royal Society of London B: Biological Sciences*, 271(1546):1407–1414, 2004.

[8] Engineering National Academies of Sciences and Medicine. *Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values*. The National Academies Press, Washington, DC, 2016.

[9] Charleston Noble, Ben Adlam, George M. Church, Kevin M. Esvelt, and Martin A. Nowak. Current CRISPR gene drive systems are likely to be highly invasive in wild populations. *eLife*, 7:–33423, 2018.

[10] Sergei V. Petrovskii, Andrew Y. Morozov, and Ezio Venturino. Allee effect makes possible patchy invasion in a predator–prey system. *Ecology Letters*, 5(3):345–352, 2002.

[11] Hidenori Tanaka, Howard A. Stone, and David R. Nelson. Spatial gene drives and pushed genetic waves. *Proceedings of the National Academy of Sciences*, page 201705868, 2017.

[12] Robert L. Unckless, Philipp W. Messer, Tim Connallon, and Andrew G. Clark. Modeling the manipulation of natural populations by the mutagenic chain reaction. *Genetics*, 201(2):425–431, 2015.

[13] Michael R. Vella, Christian E. Gunning, Alun L. Lloyd, and Fred Gould. Evaluating strategies for reversing CRISPR-Cas9 gene drives. *Scientific Reports*, 7(1):11038, 2017.

[14] Jinfeng Wang, Junping Shi, and Junjie Wei. Dynamics and pattern formation in a diffusive predator-prey system with strong allee effect in prey. *J. Differential Equations*, 251(4-5):1276–1304, 2011.

[15] Hans F. Weinberger. Invariant sets for weakly coupled parabolic and elliptic systems. *Rend. Mat. (6)*, 8:295–310, 1975. Collection of articles dedicated to Mauro Picone on the occasion of his ninetieth birthday.

[16] Bing Wu, Liqun Luo, and Xiaojing J. Gao. Cas9-triggered chain ablation of cas9 as a gene drive brake. *Nature Biotechnology*, 34:137, Feb 2016.